Shukla, Ram

ENTERED

```
RAW SEQUENCE LISTING                       DATE: 03/10/2003
PATENT APPLICATION:  US/08/984,178         TIME: 15:40:04

Input Set : N:\Crf3\RULE60\08984178.RAW.txt
Output Set: N:\CRF4\03102003\H984178.raw
```

```
         SEQUENCE LISTING
 3 (1) GENERAL INFORMATION:
 5      (i) APPLICANT: Horvitz, H. Robert
 6                     Yuan, Junying
 7                     Shaham, Shai
 9     (ii) TITLE OF INVENTION: CLONING, SEQUENCING AND CHARACTERIZATION
10                              .OF TWO CELL DEATH GENES AND USES THEREFOR
12    (iii) NUMBER OF SEQUENCES: 28
14     (iv) CORRESPONDENCE ADDRESS:
15          (A) ADDRESSEE: Fish & Richardson P.C.
16          (B) STREET: 225 Franklin Street
17          (C) CITY: Boston
18          (D) STATE: MA
19          (E) COUNTRY: USA
20          (F) ZIP: 02110-2804
22      (v) COMPUTER READABLE FORM:
23          (A) MEDIUM TYPE: Floppy disk
24          (B) COMPUTER: IBM PC compatible
25          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26          (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
28     (vi) CURRENT APPLICATION DATA:
C--> 29          (A) APPLICATION NUMBER: US/08/984,178
C--> 30          (B) FILING DATE: 03-Dec-1997
W--> 35          (C) CLASSIFICATION: 424
40    (vii) PRIOR APPLICATION DATA:
33          (A) APPLICATION NUMBER: US/08/287,669
34          (B) FILING DATE: 09-AUG-1994
37          (A) APPLICATION NUMBER: US 07/979,638
38          (B) FILING DATE: 20-NOV-1992
41          (A) APPLICATION NUMBER: US 07/897,788
42          (B) FILING DATE: 12-JUN-1992
44   (viii) ATTORNEY/AGENT INFORMATION:
45          (A) NAME: Clark, Paul T.
46          (B) REGISTRATION NUMBER: 30,162
47          (C) REFERENCE/DOCKET NUMBER: 01977/198005
49     (ix) TELECOMMUNICATION INFORMATION:
50          (A) TELEPHONE: 617/542-5070
51          (B) TELEFAX: 617/542-8906
52          (C) TELEX: 200154
54 (2) INFORMATION FOR SEQ ID NO: 1:
56      (i) SEQUENCE CHARACTERISTICS:
57          (A) LENGTH: 4407 base pairs
58          (B) TYPE: nucleic acid
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/08/984,178

DATE: 03/10/2003
TIME: 15:40:04

Input Set : N:\Crf3\RULE60\08984178.RAW.txt
Output Set: N:\CRF4\03102003\H984178.raw

```
     59          (C) STRANDEDNESS: single
     60          (D) TOPOLOGY: linear
W--> 62     (ii) MOLECULE TYPE: DNA
     64     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
     66 GAATTCGCGT CGAATCATTG TCTGTTCGGT ATCGATTCAG AAACCGAAAC TTGTGATCGA    60
     68 TAACAAGTCA TTCAAACACG GCGAAGATGT CTATGCGTAT AACAGAATAT TTGGAGAAAT   120
     70 GCTCGCAAAA CTCGAAATTG TCACCGATAA AATGATTAAC TTGAAGGGGC TAATGTAAGT   180
     72 TATCTGATGT TTCTACAATT AAAAAAATTG TTTTTTTTTC CAAATTAATT TTCGAAGATT   240
     74 AACGAAAAAC GATTAAAAAT CAATAAAACG CAATAAAGAG GGCTTGGCTT TCTTTTTAAT   300
     76 TTAAATTATA ATTTTTCTGA TTGTTGTATG AAGCTACAAA ATGTACTGTT TTTGTATTTG   360
     78 AATATTGTAT TACAGGGTTG GGATTCTCGG CAAATATCAG CGACAGTGGA AGATTTAGAA   420
     80 GAAGGACGTG TGACAATCAC TAAGTCAAAG AGGGAAAGGA TAAAGGATTG TGATATTTCA   480
     82 CTGTTTTACT CATTCGCTTT TTAAATAAGA ACTATATGCC GATTGCCGA TATATTTTTG   540
     84 TTTATTAGGC CTCTCACATT CCTGTACAAT GTTCTACCA AATAAACTGC ATTTTTATCT   600
     86 GAAAATTCGA ATTTATTTTT GTCTACTTTT TACTCGTTGC ATTCGAGATC AGCATATCTT   660
     88 CCGGTCTATT TATATTCAAC GATTTTTATA AATTAGTACT CCTTCATGTT TAATTTCATT   720
     90 TTATCTGTAA GCTTTACTGT ATTTTTTTAA AATCTTTCTT GCTTCTATCT GATTATACAA   780
     92 TGTTCTTTAC TCATTTTCAA GGTATTTTTA TGCCTCACAA TTTATGCACA TTTCGGGCTT   840
     94 GGAGATTTAT CCTCTATATT ACATGCCTGT TTTTTAAAG GATATAATGT TTAACAAATA   900
     96 ATTTTTTATC AATGCTATTG TATATTCTCC AGCTAACCGT TGTTTCGAAA ACATCACCTA   960
     98 GCATTTTAAA ATTCACAAAA TCTTGCTTCC TTATAATCAA GAAGATTTTT CAGATGCTCT  1020
    100 GCGAAATCGA ATGCCGCGCT TTGAGCACGG CACACACGAG GCTCATCCAC GACTTTGAAC  1080
    102 CACGTGACGC ATTGACTTAT TTAGAAGGCA AAAACATTTT CACAGAAGAT CATTCTGAAC  1140
    104 TTATCAGTAA AATGTCAACT CGCCTCGAGA GGATCGCCAA TTTTCTTCGA ATCTATCGAC  1200
    106 GTCAAGCTTC TGAACTTGGA CCACTCATCG ACTTTTTCAA CTACAACAAT CAAAGTCACC  1260
    108 TTGCTGATTT CCTCGAAGAC TACATCGATT TTGCGATAAA TGAGCCAGAT CTACTTCGTC  1320
    110 CAGTAGTGAT TGCTCCACAA TTTTCCCGAC AAATGCTCGA TAGGAAACTA TTGCTTGGGA  1380
    112 ATGTTCCAAA ACAAATGACA TGCTATATTC GAGAGTATCA CGTGGATCGA GTGATCAAAA  1440
    114 AGCTCGACGA GATGTGTGAT TTAGGTGAGA AAACTGGAAG CTCTCGTGTT TATTATAATC  1500
    116 TTGCTTAAAC TTCAGACTCC TTTTTTCTGT TTCTACACGG CCGAGCTGGA TCCGGAAAAT  1560
    118 CAGTAATTGC ATCACAAGCT CTTTCGAAAT CTGACCAACT TATTGGAATG TGAGTGGTAT  1620
    120 TATCTGAATC TACGGATCTT CATTCTATTA CAGAAATTAT GATTCAATCG TTTGGCTCAA  1680
    122 AGATAGTGGA ACAGCTCCAA AATCTACATT CGATTTATTT ACGGATATTT TGCTGATGCT  1740
    124 AAAGTGAGTG AATAGAGTGC ATGTAACATT CAGCATGATT TTGAAATTAT GAAAATTTGA  1800
    126 CCTGGTTAGC TTTTAATTTG ATATTTCGTG ACGCTTGCAT GTTTTGTGTG TTTGAAGACG  1860
    128 AGCCCGTGTT GTGAGCGACA CGGATGACTC GCATTCGATC ACCGACTTCA TTAACCGTGT  1920
    130 TCTTTCAAGA AGCGAAGACG ATCTTCTCAA TTTCCCATCG GTGGAGCATG TCACGTCAGT  1980
    132 TGTACTCAAA AGGATGGTAA GTTGCTTGCC GATTCTGGTA CAATATCTTA AATTATTGGT  2040
    134 TTTTAGATCT GCAACGCACT CATTGATCGT CCAAATACTT TATTCGTATT TGATGACGTA  2100
    136 GTTCAAGAAG AAACAATTCG TTGGGCTCAG GAGCTACGTC TTCGATGTCT TGTAACTACT  2160
    138 CGTGACGTGG AAATATCAAA TGCTGCTTCT CAAACATGCG AATTCATTGA AGTGACATCA  2220
    140 TTGGAAATCG ATGAATGTTA TGATTTTCTA GAAGCTTATG GAATGCCGAT GCCTGTTGGA  2280
    142 GAAAAGAAG AAGATGTGCT TAATAAAACA ATCGAACTAA GCAGTGGAAA TCCAGCAACG  2340
    144 CTTATGATGT TTTTCAAGTC TTGTGAACCG AAAACATTTG AAAAGTGAGT GGGACATACC  2400
    146 AATTTGAGAC TTTTAAAATA ATTTATTCTA CAATAAAAGT TAATCAAAAA GTTTCATAGC  2460
    148 TGATTGTCTT TAAATTTTAC GAATTGAGGA TCAAAATCAA GAATTAGGAT CCTGGCACGA  2520
    150 GAGAAACTG TGTAGCTACC GTACCCGAGA GATTTTCTTG ATATTTGCCA TCGATTTAAT  2580
    152 TTTTTAAGAA AATTATCGTT TTACATAATT GAACAAGAGA TACACGGTCT CGACCCGACG  2640
    154 GAAATTTTTT AAATGAAAGC GAGTATGAGC CTGTTTTCAT TATTTTTCGA TTTTCTCTTG  2700
```

RAW SEQUENCE LISTING   DATE: 03/10/2003
PATENT APPLICATION: US/08/984,178   TIME: 15:40:04

Input Set : N:\Crf3\RULE60\08984178.RAW.txt
Output Set: N:\CRF4\03102003\H984178.raw

```
156 TTGTTTCTTT TTATTTAAAG CCTTTTATTT TGAAACAAGT CTAAAAATAT TAAAAACTGA   2760
158 ATAAAATATT TAAAAAAAAT CAAGTAAAAT AGAAAAACAG CAAGGCTGGA GACTACTGTA   2820
160 CTTCTTAAAT CCGCATACTC TTTTTATTTA ATCATTTTCC GGAATGTCGA AACGAAATAA   2880
162 TACATTTTTA GTCCAAAATC GCTAGGTATA TTCTTAAAAT TATCAAACAT TTTGCATTCA   2940
164 GAATGGCACA GCTTAATAAC AAAATTGGAAA GTCGAGGATT AGTCGGTGTT GAATGTATCA   3000
166 CCCCTTACTC GTACAAGTCA CTCGCAATGG CTCTTCAAAG ATGTGTTGAA GTTTTGTCAG   3060
168 ATGAGGATCG AAGTGCTCTT GCTTTCGCAG TTGTGATGCC TCCTGGAGTT GATATACCCG   3120
170 TCAAGCTATG GTCATGTGTT ATTCCAGTTG ATATTTGTTC AAATGAAGAA GAACAATTGG   3180
172 ATGATGAAGT TGCGGATCGG TTGAAAAGAC TCAGCAAGTA TGAGTCTTGA AATTTGAAGA   3240
174 TTTAAATTAA CACTTAAAAT TTCAGACGTG GAGCTCTTCT CAGTGGAAAA CGAATGCCCG   3300
176 TTTTGACATT CAAAATTGAT CATATTATCC ATATGTTCTT GAAACACGTC GTTGATGCAC   3360
178 AAACTATCGC CGTATGCTGA AAATGTCTCA ACTTTCAATT AAATTTTAAA TTTTCAGAAT   3420
180 GGAATCTCAA TTCTCGAGCA GCGTCTTCTT GAAATAGGAA ACAATAATGT ATCAGTACCG   3480
182 GAGCGACATA TACCATCACA TTTCCAAAAA TTCCGTCGTT CATCAGCCAG TGAGATGTAT   3540
184 CCAAAAACTA CAGAAGAAAC TGTGATCCGT CCTGAAGACT TCCCAAAGTT CATGCAATTG   3600
186 CACCAGAAAT TCTATGACTC CCTCAAAAAT TTTGCATGCT GTTAAAACCT ATCGTGTACA   3660
188 ATATTGCCTG TATATTCCCC TCGAAATACG TTTATACTTT TTCGCACGAG TTTTCTCATT   3720
190 TTTTCATTTG TACTTGTTTT ATTTCTCTCC AAAATTTCAG ATCTATCCCA AATGTTCTTA   3780
192 AATTTAATGT TTTCTACAGA TACTCAACAC ATCTTGTTTC ATCTCATCCT TGCTTTTTTT   3840
194 TTTCAAATAT ATTCAGTTTC TTTTATAATT TTAATTAATC GAATTAATAC ATTCACGTAA   3900
196 AGAATTTCGT GGACTATTAT TTTATCGCAT CCAAATGATT TATTCCCTAT TGTTCGAAAC   3960
198 TTCCAAATTG ATCATTTTTA AACACGCCTC ATTAAATTGA AAGTCGTACT TTTAGTCTCG   4020
200 AACATGAAGT AAGTTATTTT CTGTGTTCTA AATTCAAAGT GCATTCCAAA AGGACATTTG   4080
202 ATGAGTTTTC ACGAAAACCG TAATTTTTAC AATTTCCTTT CAGTTTTGAA GATGTTCGAT   4140
204 TTCTTTCCTC TGTTGGCGTC ATTACTACAT TTGCTTTGCT GCTTCACTTT ATCGAGATTC   4200
206 TTGCCATCAA TGGAGTTCCA TCTAGACCGA TAGCAGTCTT CATATCATTA TCCCTGTATA   4260
208 TTGTACTGTT TCAGTATTTT AACTTATCGA TTACGTACTA TATTCAGTGG TTCACTGTTT   4320
210 TCGGTCAATG GGTGACACGT GCTCGACGAN NAATTTTCAA CGAACGCAAT CTCCTAGTCA   4380
212 CTTATCAACC AAGAGCCCTC ACCCATG                                       4407
```

214 (2) INFORMATION FOR SEQ ID NO: 2:
216   (i) SEQUENCE CHARACTERISTICS:
217   (A) LENGTH: 549 amino acids
218   (B) TYPE: amino acid
219   (C) STRANDEDNESS: Not Relevant
220   (D) TOPOLOGY: linear
222   (ii) MOLECULE TYPE: protein
224   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
226   Met Leu Cys Glu Ile Glu Cys Arg Ala Leu Ser Thr Ala His Thr Arg
227   1   5   10   15
229   Leu Ile His Asp Phe Glu Pro Arg Asp Ala Leu Thr Tyr Leu Glu Gly
230   20   25   30
232   Lys Asn Ile Phe Thr Glu Asp His Ser Glu Leu Ile Ser Lys Met Ser
233   35   40   45
235   Thr Arg Leu Glu Arg Ile Ala Asn Phe Leu Arg Ile Tyr Arg Arg Gln
236   50   55   60
238   Ala Ser Glu Leu Gly Pro Leu Ile Asp Phe Phe Asn Tyr Asn Asn Gln
239   65   70   75   80
241   Ser His Leu Ala Asp Phe Leu Glu Asp Tyr Ile Asp Phe Ala Ile Asn
242   85   90   95

RAW SEQUENCE LISTING
PATENT APPLICATION: US/08/984,178

DATE: 03/10/2003
TIME: 15:40:04

Input Set : N:\Crf3\RULE60\08984178.RAW.txt
Output Set: N:\CRF4\03102003\H984178.raw

```
244    Glu Pro Asp Leu Leu Arg Pro Val Val Ile Ala Pro Gln Phe Ser Arg
245                100                 105                 110
247    Gln Met Leu Asp Arg Lys Leu Leu Leu Gly Asn Val Pro Lys Gln Met
248        115                 120                     125
250    Thr Cys Tyr Ile Arg Glu Tyr His Val Asp Arg Val Ile Lys Lys Leu
251        130                 135                 140
253    Asp Glu Met Cys Asp Leu Asp Ser Phe Phe Leu Phe Leu His Gly Arg
254    145             150                 155                     160
256    Ala Gly Ser Gly Lys Ser Val Ile Ala Ser Gln Ala Leu Ser Lys Ser
257                165                 170                 175
259    Asp Gln Leu Ile Gly Ile Asn Tyr Asp Ser Ile Val Trp Leu Lys Asp
260                180                 185                 190
262    Ser Gly Thr Ala Pro Lys Ser Thr Phe Asp Leu Phe Thr Asp Ile Leu
263        195                 200                 205
265    Leu Met Leu Lys Ser Glu Asp Asp Leu Leu Asn Phe Pro Ser Val Glu
266        210                 215                 220
268    His Val Thr Ser Val Val Leu Lys Arg Met Ile Cys Asn Ala Leu Ile
269    225                 230                 235                 240
271    Asp Arg Pro Asn Thr Leu Phe Val Phe Asp Asp Val Val Gln Glu Glu
272        275                 245                 250             255
274    Thr Ile Arg Trp Ala Gln Glu Leu Arg Leu Arg Cys Leu Val Thr Thr
275                260                 265                 270
277    Arg Asp Val Glu Ile Ser Asn Ala Ala Ser Gln Thr Cys Glu Phe Ile
278        275                 280                 285
280    Glu Val Thr Ser Leu Glu Ile Asp Glu Cys Tyr Asp Phe Leu Glu Ala
281        290                 295                 300
283    Tyr Gly Met Pro Met Pro Val Gly Glu Lys Glu Glu Asp Val Leu Asn
284    305                 310                 315                 320
286    Lys Thr Ile Glu Leu Ser Ser Gly Asn Pro Ala Thr Leu Met Met Phe
287                325                 330                 335
289    Phe Lys Ser Cys Glu Pro Lys Thr Phe Glu Lys Met Ala Gln Leu Asn
290                340                 345                 350
292    Asn Lys Leu Glu Ser Arg Gly Leu Val Gly Val Glu Cys Ile Thr Pro
293        355                 360                 365
295    Tyr Ser Tyr Lys Ser Leu Ala Met Ala Leu Gln Arg Cys Val Glu Val
296        370                 375                 380
298    Leu Ser Asp Glu Asp Arg Ser Ala Leu Ala Phe Ala Val Val Met Pro
299    385                 390                 395                 400
301    Pro Gly Val Asp Ile Pro Val Lys Leu Trp Ser Cys Val Ile Pro Val
302                405                 410                 415
304    Asp Ile Cys Ser Asn Glu Glu Glu Gln Leu Asp Asp Glu Val Ala Asp
305                420                 425                 430
307    Arg Leu Lys Arg Leu Ser Lys Arg Gly Ala Leu Leu Ser Gly Lys Arg
308        435                 440                 445
310    Met Pro Val Leu Thr Phe Lys Ile Asp His Ile Ile His Met Phe Leu
311        450                 455                 460
313    Lys His Val Val Asp Ala Gln Thr Ile Ala Asn Gly Ile Ser Ile Leu
314    465                 470                 475                 480
316    Glu Gln Arg Leu Leu Glu Ile Gly Asn Asn Asn Val Ser Val Pro Glu
```

```
317                      485              490                   495
319      Arg His Ile Pro Ser His Phe Gln Lys Phe Arg Arg Ser Ser Ala Ser
320                  500              505              510
322      Glu Met Tyr Pro Lys Thr Thr Glu Glu Thr Val Ile Arg Pro Glu Asp
323              515              520              525
325      Phe Pro Lys Phe Met Gln Leu His Gln Lys Phe Tyr Asp Ser Leu Lys
326          530              535              540
328      Asn Phe Ala Cys Cys
329      545
331 (2) INFORMATION FOR SEQ ID NO: 3:
333      (i) SEQUENCE CHARACTERISTICS:
334          (A) LENGTH: 12 amino acids
335          (B) TYPE: amino acid
336          (C) STRANDEDNESS: Not Relevant
337          (D) TOPOLOGY: linear
339      (ii) MOLECULE TYPE: protein
341      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
343      Asp Gln Asp Lys Ser Gly Phe Ile Glu Glu Asp Glu
344      1               5                   10
346 (2) INFORMATION FOR SEQ ID NO: 4:
348      (i) SEQUENCE CHARACTERISTICS:
349          (A) LENGTH: 12 amino acids
350          (B) TYPE: amino acid
351          (C) STRANDEDNESS: Not Relevant
352          (D) TOPOLOGY: linear
354      (ii) MOLECULE TYPE: protein
356      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
358      Asp Gln Asp Lys Asp Asp Phe Ile Gly Glu Asp Glu
359      1               5                   10
361 (2) INFORMATION FOR SEQ ID NO: 5:
363      (i) SEQUENCE CHARACTERISTICS:
364          (A) LENGTH: 12 amino acids
365          (B) TYPE: amino acid
366          (C) STRANDEDNESS: Not Relevant
367          (D) TOPOLOGY: linear
369      (ii) MOLECULE TYPE: protein
371      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
373      Asp Ser Asp Gly Asp His Lys Ile Gly Val Asp Glu
374      1               5                   10
376 (2) INFORMATION FOR SEQ ID NO: 6:
378      (i) SEQUENCE CHARACTERISTICS:
379          (A) LENGTH: 12 amino acids
380          (B) TYPE: amino acid
381          (C) STRANDEDNESS: Not Relevant
382          (D) TOPOLOGY: linear
384      (ii) MOLECULE TYPE: protein
386      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
388      Asp Ile Asn Lys Asp Asp Val Val Ser Trp Glu Glu
389      1               5                   10
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 03/10/2003
PATENT APPLICATION: US/08/984,178   TIME: 15:40:05

Input Set : **N:\Crf3\RULE60\08984178.RAW.txt**
Output Set: **N:\CRF4\03102003\H984178.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 4350,4351
Seq#:20; Xaa Pos.94,95,96,120,179,318
Seq#:21; Xaa Pos.310

VERIFICATION SUMMARY
PATENT APPLICATION: **US/08/984,178**

DATE: 03/10/2003
TIME: 15:40:05

Input Set : **N:\Crf3\RULE60\08984178.RAW.txt**
Output Set: **N:\CRF4\03102003\H984178.raw**

```
L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:35 M:238 W: Alpha Fields not Ordered, Reordered [(C) CLASSIFICATION:] of (1)(vi)
L:62 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1
L:457 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=11
L:564 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=18
L:959 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20 after pos.:80
M:341 Repeated in SeqNo=20
L:1109 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21 after pos.:304
L:1150 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=22
L:1164 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=23
L:1178 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=24
L:1192 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=25
L:1206 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=26
L:1220 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=27
L:1234 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=28
```